

日本国特許庁

PATENT OFFICE  
JAPANESE GOVERNMENT

1177, 743  
30307-168819  
2062

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出願年月日

Date of Application:

2000年12月14日

出願番号

Application Number:

特願2000-380781

出願人

Applicant(s):

経済産業省産業技術総合研究所長

後藤 真孝

伊藤 克亘

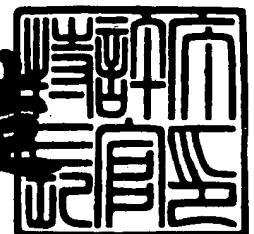


CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 3月16日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2001-3012760

【書類名】 特許願

【整理番号】 10744210

【提出日】 平成12年 7月13日

【あて先】 特許庁長官 殿

【国際特許分類】 G10L 9/00

【発明者】

【住所又は居所】 茨城県つくば市梅園1丁目1番4 工業技術院電子技術  
総合研究所内

【氏名】 後藤 真孝

【発明者】

【住所又は居所】 茨城県つくば市梅園1丁目1番4 工業技術院電子技術  
総合研究所内

【氏名】 伊藤 克亘

【特許出願人】

【代表出願人】

【識別番号】 000001144

【氏名又は名称】 工業技術院長 梶村 皓二

【特許出願人】

【識別番号】 592163860

【氏名又は名称】 後藤 真孝

【特許出願人】

【住所又は居所】 茨城県つくば市梅園1丁目1番4 工業技術院電子技術  
総合研究所内

【氏名又は名称】 伊藤 克亘

【指定代理人】

【識別番号】 220000356

【氏名又は名称】 工業技術院電子技術総合研究所長 児玉 皓雄

【電話番号】 0298-61-2175

【代理関係の特記事項】 特許出願人 工業技術院長の指定代理人

【その他】 国以外の全ての者の持分の割合 50 / 100

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声補完入力方法及び装置

【特許請求の範囲】

【請求項 1】 ユーザの発声をシステムが認識するものにおいて、ユーザの発する言語の断片に対し、システムが残りの部分を補う音声補完入力方法。

【請求項 2】 請求項 1 における音声補完入力方法において、ユーザの意図的な有声休止をトリガーとして、システムが残りの部分を補う音声補完入力方法。

【請求項 3】 請求項 1 における音声補完入力方法において、ユーザの意図的な有声休止を含む特定のキーワードの発声により、該キーワードを任意の文字列とみなし、システムがその前後の関係より前記キーワード部分を置き換える音声補完入力方法。

【請求項 4】 請求項 1 ないし 3 のいずれかに記載された音声補完入力方法において、補完候補が複数ある場合には、システムが補完候補一覧を画面に表示又は音声で読み上げ、ユーザが音声、キーボード又はその他のデバイスを用いて選択し、補完候補が一つの場合には、システムがユーザに画面表示や音声合成等により確認を求めるか、あるいは、自動的に最後まで入力する音声補完入力方法。

【請求項 5】 ユーザの発声をシステムが認識するものにおいて、ユーザの発する言語の断片に対し、システムが残りの部分を補完する音声補完入力装置。

【請求項 6】 請求項 5 における音声補完入力装置において、ユーザの意図的な有声休止をトリガーとして、システムが残りの部分を補う音声補完入力装置。

【請求項 7】 請求項 5 における音声補完入力装置において、ユーザの意図的な有声休止を含む特定のキーワードの発声により、該キーワードを任意の文字列とみなし、システムがその前後の関係より前記キーワード部分を置き換える音声補完入力装置。

【請求項 8】 請求項 5 ないし 7 のいずれかに記載された音声補完入力装置において、補完候補が複数ある場合には、システムが補完候補一覧を画面に表示

又は音声で読み上げ、ユーザが音声、キーボード又はその他のデバイスを用いて選択し、補完候補が一つの場合には、システムがユーザに画面表示や音声合成等により確認を求めるか、あるいは、自動的に最後まで入力する音声補完入力装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】

本願発明は、音声認識において一般的に適用される音声補完入力方法及び装置に関するものである。

【0002】

【従来の方法】

人と人が交わす音声対話が、我々人間にとって快適で優れた情報交換手段である一つの理由は、対話相手が様々な形で自分の発話や思考の手助けをしてくれるからである。すなわち、人間同士の音声対話においては、たとえ話者がある単語を最後まで思い出せず、その途中に言い淀むことがあったとしても、対話相手は、話者の言いたいことを推測して候補を提示することで、話者が思い出すのを手助けしてくれるのである。例えば、話者が「音声補完」という単語全体を思い出せず「おんせー」と言い淀むと、対話相手が「音声補完？」と問いかけて手助けをしてくれる。これは、発話された単語の断片の続きを補うことで、話者が述べようとしている単語全体の候補を提示している、つまり、単語を補完していると見なすことができる。

【0003】

補完の概念は、テキストインタフェースにおいては既に広く受け入れられている。例えば、tcshやbash等のUNIXシェル及びEmacs/Mule等のテキストエディタは、ファイル名やコマンド名の補完機能を提供している。こうした補完機能においては、ユーザが補完機能と呼び出すキー（以下「補完トリガーキー」という。）を押したときに、途中までタイプされた単語の断片の続きが補われる。他にも、Netscape Communicator及びInternet Explorer等のWWWブラウザにおいては、URL等の自動補完機能が導入されている。これは、ユーザがタイプしている最中に、シ

ステム側が補完候補一覧を次々と提示していく機能である。

【0004】

また、補完機能は、近年、テキストインタフェース以外でも導入され始めている。例えばペン入力においては、予測ペン入力インタフェースやPOBox等の自動補完機能を持ったインタフェースが提案されている（予測ペン入力に関しては、福島俊一、山田洋志「予測入力インタフェースとその手書き操作削減効果」情報処理学論、Vol.37, No.1, pp.23-30(1996)、POBoxについては、Masui, T. 「An Efficient Text Input Method for Pen-based Computers」 Proceedings of the ACM Conference on Human Factors in Computing Systems(CHI'98), pp.328-335(1998)参照）。

【0005】

【発明が解決しようとする課題】

しかし、音声入力インタフェースにおいては、音声入力中に補完機能呼び出す適当な手段がなかったために、音声補完入力が存在しないというのが実状であった。

【0006】

【課題を解決するための手段】

本願発明は、システムに対する音声入力中に、ユーザが入力したいすべてを発話せずにその断片だけを発話しても、システムが残りを補って入力することを可能にする音声入力インタフェース機能（以下「音声補完」という。）を導入することにより、音声認識を中心とした音声インタフェースをより使いやすくすることを可能とした。

【0007】

音声補完の方式には、テキスト補完と同様に、補完トリガーキーによる補完と、発声中に次々と補完候補を出すような自動補完の二つが考えられる。しかし、音声で自動補完をしようとしても、発話した断片を認識する時点で既に曖昧性が大きいため、テキスト補完と同じような精度で適切な候補を提示し続けることはできず、自動補完機能自体が煩わしくなってしまう可能性が高い。つまり、音声の自動補完は不適切な機能となりやすい。そこで、音声補完の場合には、ユーザ

が候補を見たいと思うタイミングで、自発的に補完機能呼び出せることが重要となる。使いやすい音声補完を実現するためには、どのように補完機能呼び出すか、つまり音声における補完トリガーキーを何にするかが鍵となる。

## 【 0 0 0 8 】

本願発明においては、言い淀み現象の一つである有声休止(filled pause)に補完トリガーキーの役割を担わせることで、ユーザが自分の意志で、労力をかけずに補完機能呼び出すことを可能にする。ここで有声休止とは、つなぎ語と音節の引き延ばし（日本語では母音の引き延ばし）の二つに分類され、つなぎ語とは、「えー」、「うー(ん)」、「あー」、「まー」、「んー」、「あの一」、「その一」、「この一」のようなものである。補完トリガーキーとして、音声入力中に有声休止を行うことは、人間にとって極めて自然な行為である。

実際、人間同士の対話でも、有声休止は似た役割で用いられることがあり、話者は、次の単語を思い出すまでの間、時間を稼ぐために有声休止を用いたり、場合によっては対話相手の助けを期待しながら有声休止を用いたりしているのである。

## 【 0 0 0 9 】

次に、音声補完入力の2つの様態について説明する。

## [有声休止を用いた音声補完方法]

第1の方法は、言い淀み現象の一つである有声休止に、補完機能呼び出す補完トリガーの役割を担わせることで、ユーザが自分の意志ではあるが、あまり労力をかけずに音声補完を行うことを可能にする。

例えば、「音声補完」が音声認識システムの辞書に登録されている場合には、「おんせー」と「い」の引き延ばしにより有声休止を行うか、あるいは、「おんせい、えー」とつなぎ語「えー」により有声休止を行うと、「音声補完」のように補完されて表示される。補完候補が複数ある場合には、システムがその候補を画面に表示あるいは音声合成をしたりして提示し、ユーザがその中から適切な候補を選べばよい。補完候補が一つの際には、ユーザに確認を求めてもよいし、自動的に最後まで入力してもよい。

## 【 0 0 1 0 】

## [音声ワイルドカード補完方法]

第2の方法は、ユーザが発声途中で意図的に有声休止を行いながらある特定のキーワード(「なんかー」等)を発声すると、そのキーワード全体をワイルドカード(任意の文字列)とみなし、システムがその前後からワイルドカード部分を補って入力することを可能にする。

例えば、「音声補完」が音声認識システムの辞書に登録されているときに、「なんかー」という有声休止を伴うキーワードを使って、「なんかーほかん」と発声すれば、「なんかー」の部分を任意の文字列に置き換えた候補一覧(「音声補完」「音声ワイルドカード補完」等)を見て選択できる。補完候補が複数ある場合には、システムが画面に表示したり音声合成をしたりして提示し、ユーザがその中から適切な候補を選ぶ。補完候補が一つの場合には、ユーザに確認を求めてもよいし、自動的に最後まで入力してもよい。

## 【0011】

## 【実施例】

音声補完は、単語や文節、文章等様々なレベル対象として実施が可能であるが、以下、単語を例にとって、有声休止を用いた音声補完について説明する。ただし、ここでの単語は、音声認識システムの単語辞書上(言語モデル上)の1単語とする。したがって、例えば「宇多田ヒカル」のような姓名が一つの単語として登録されているときに、「うただー」と「だ」の音で有声休止を行うと、「宇多田ヒカル」が補完候補の一つとして得られる。

ユーザは、以下のように有声休止を用いて音声補完しながら、単語を入力することができる(図1参照)。

## 【0012】

1. 単語の発声途中で母音を引き延ばすと、既に発声された断片から始まる補完候補(単語)の一覧が、番号付きで即座に表示される。

例えば、「うただー」と入力すると(図2-1)(図2-2)、

「1. 宇多田ヒカル、 2. 上原多香子、 3. MR.DYNAMITE」

のように補完候補が表示される(図2-3)。

## 【0013】



2. 補完候補が多くて画面に入りきらない場合には、「次の候補」というマークが表示される。この場合、「次」「次の」「次の候補」のいずれかを発声すると他の候補を見ることができる。補完候補が不適切なとき、あるいは別の単語を入力したくなったときは、次の3.の選択をせずに別の発話に移ってもよい。

【0014】

3. ユーザは候補一覧を見ながら、以下の3通りの方法で補完候補を選択できる。

(a) 候補の番号を言って選択する(図2-4)。

(例えば、「1番」か「1」と発声する。)

(b) 単語の続きを読み上げて選択する。

(例えば、「ひかる」と発声する。)

(c) 単語全体を頭から読み上げて選択する。

(例えば、「うただひかる」と発声する。)

(d) 候補を他のデバイス(キーボードやマウス、タッチパネル等)で選択する。

選択すると、その候補は強調表示され(図2-5)、音声認識結果として確定される(図2-6)。

【0015】

音声補完は、一つの単語を入力中に、繰り返し呼び出すことが可能である。例えば、「サザンオールスターズ」を入力するときに、「さざんー」で候補一覧を見た後、「おーるー」でさらに絞り込まれた候補一覧を見て、最後に「すたーず」と言って確定できる。なお、この例に示したように、単語中の長母音(「おーる」の /o-/)においては音声補完が呼び出されず、意図的に有声休止した箇所でのみ呼び出されるようにする必要がある。

【0016】

実際に、音声補完が可能な音声入力インタフェースシステムを実装し、運用した結果、音声補完が実用的に機能し、ユーザが音声補完を呼び出しながら、インタラクティブに音声入力できることを確認した。音声補完は、使用するのが容易で訓練は不要であり、直感的で使いやすいインタフェース機能であることがわかった。特に、長い語句を入力する際に、音声補完は非常に有効であった。

今回の運用では、曲名やアーティスト名の入力、住所入力といった各種固有名詞の入力で有用性を確認したが、システムに対して音声で入力するような様々な局面にもすぐに適用できるものである。

#### 【0017】

##### 【発明の効果】

まず第1に、記憶の補助がある。たとえ、入力したい内容がうろ覚えであっても、部分的に思い出して発声することができれば、システムの手助けにより入力することが可能になることである。

次に、入力したい内容が長くて複雑なときにおいても、その内容を特定するのに十分な部分だけを発声することにより、システムが残りを補完して入力することができるという効果がある。

また、従来の音声インタフェースの多くが、ユーザ側にすべての音を最後まで丁寧に発声することを強いていたのに対し、音声補完においては、単語や文節、文章等の断片を発声することにより入力できるため、心理的抵抗が少なく使いやすいという効果がある。

#### 【0018】

##### 【図面の簡単な説明】

【図1】音声補完の操作の流れ

【図2-1】「うただー」と入力中の画面

【図2-2】「だー」の有声休止中の画面

【図2-3】補完候補表示画面

【図2-4】「1番」と入力した直後の画面

【図2-5】1番の候補が輝いている画面

【図2-6】1番の候補が確定した画面

【書類名】図面

【図1】

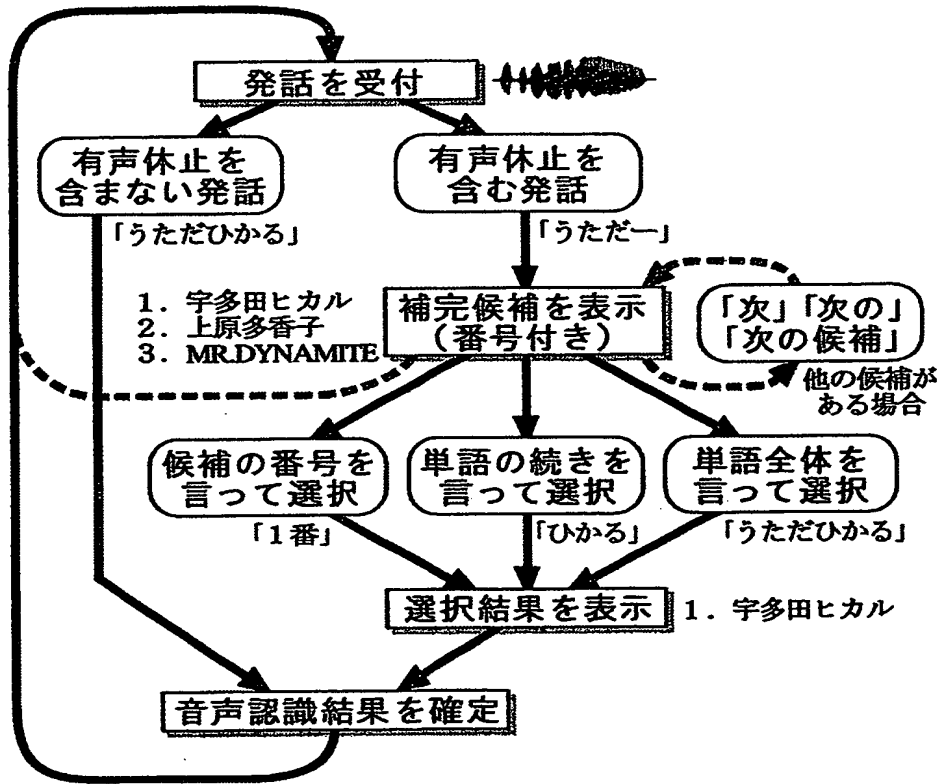


図 1: 音声補完の操作の流れ

【図 2 - 1】

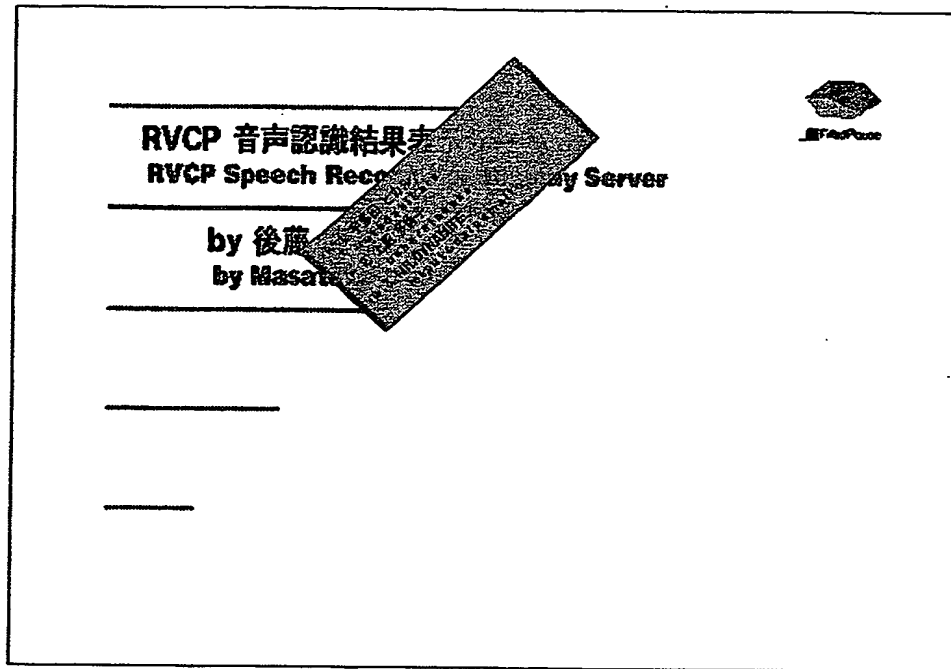


図 2 - 1 : 「うただー」と入力中の画面

【図 2-2】

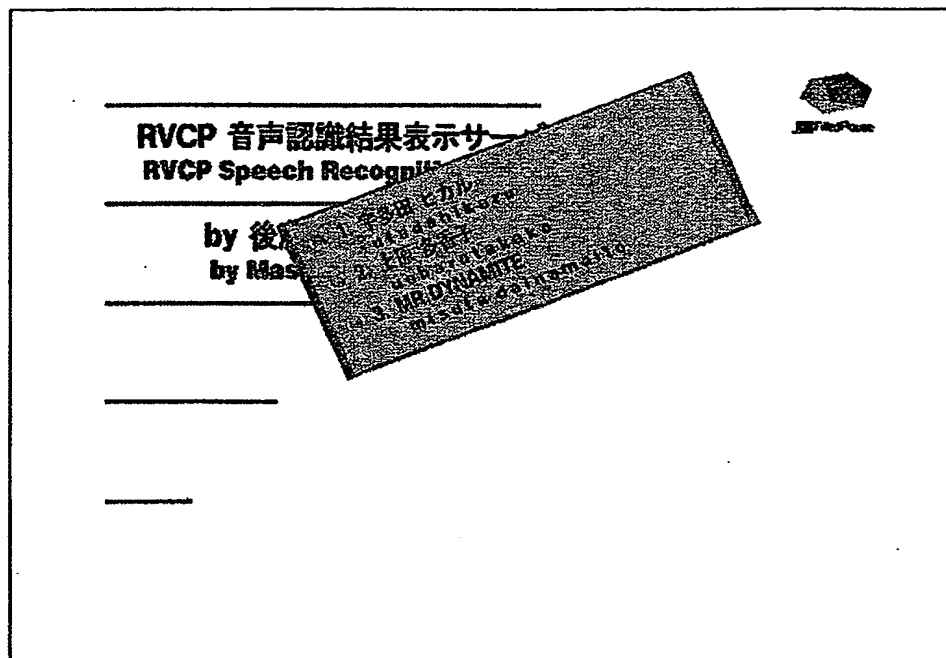


図 2-2 : 「だー」 の有声休止中の画面

【図 2-3】

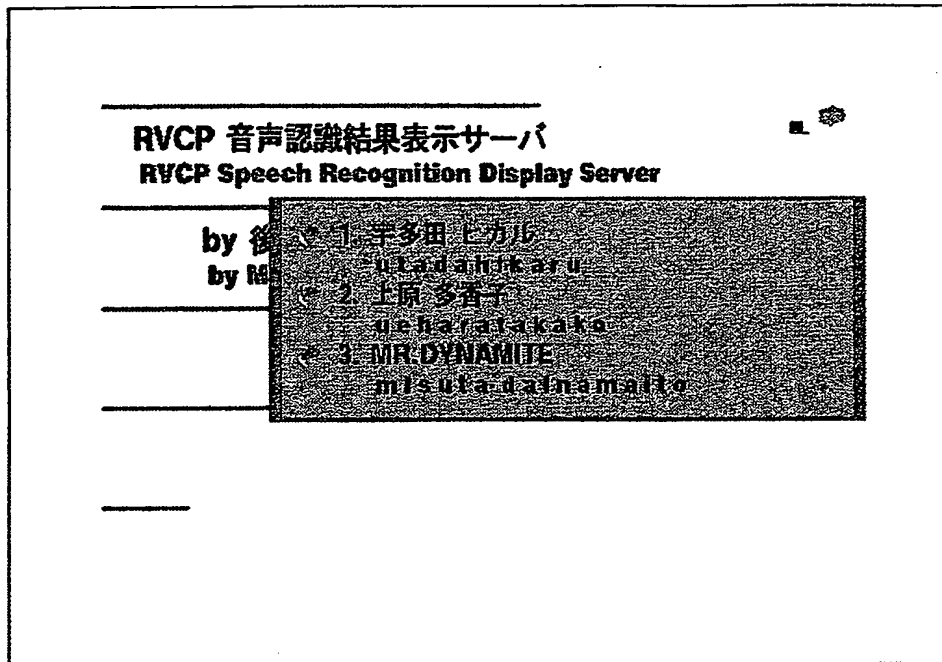


図 2-3 : 補完候補表示画面

【図2-4】

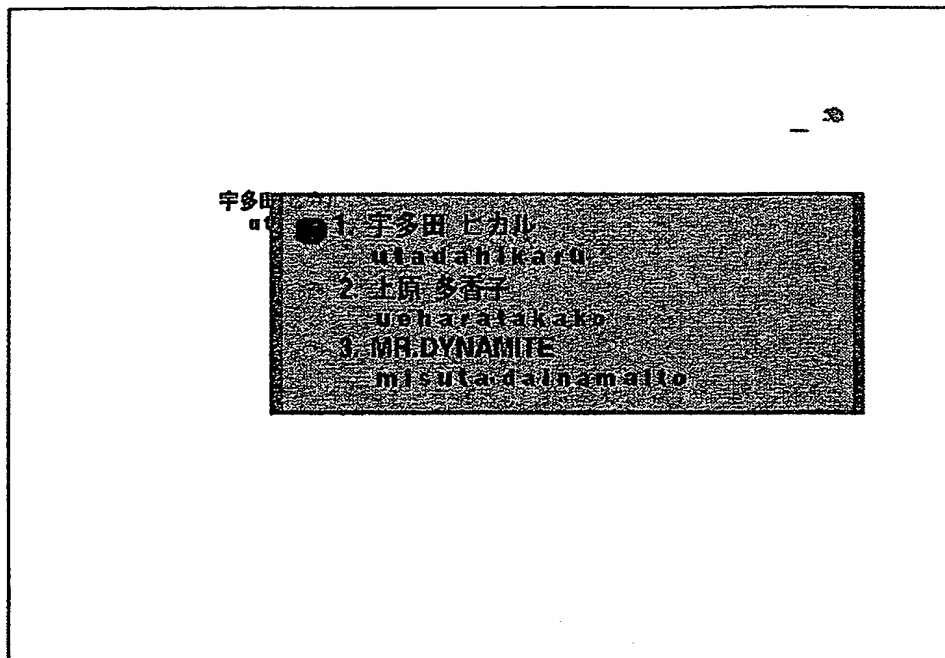


図2-4 : 「1番」と入力した直後の画面

【図2-5】

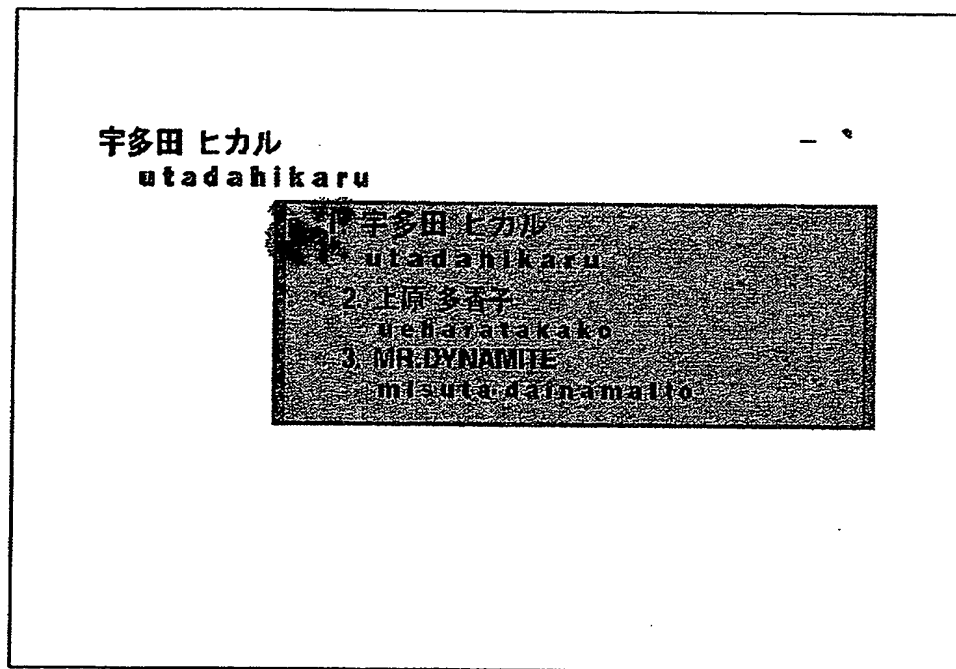


図2-5：1番の候補が輝いている画面



【図 2 - 6】

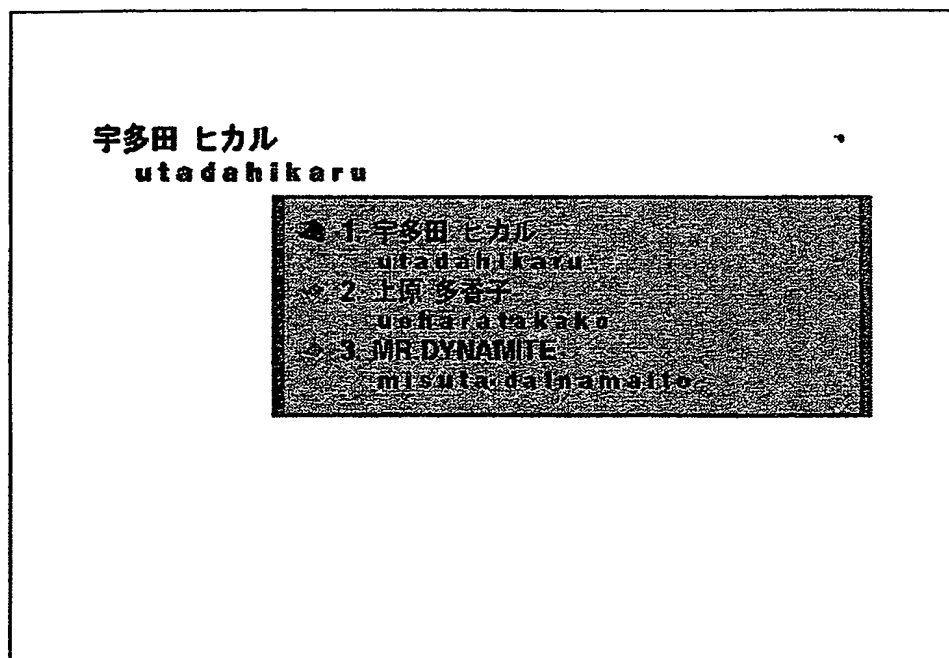


図 2 - 6 : 1 番の候補が確定した画面

【書類名】 要約書

【要約】

【課題】 一般の音声認識システムによる音声入力では、入力したい内容を正確にすべて発声しなければ入力できず、うろ覚えの単語や文章は入力できなかった。また、入力したい内容が長くて複雑なときに、人間同士ならば、その内容を特定するのに十分な部分まで発声すれば伝わるのに、音声認識システムに対しては最後まですべて発声する必要があった。

【解決手段】 有声休止をトリガーとして採用することにより音声補完を可能とした。

【選択図】 図1

認定・付加情報

特許出願の番号	特願 2000-212158
受付番号	50000882232
書類名	特許願
担当官	宇留間 久雄 7277
作成日	平成 12 年 10 月 20 日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000001144
【住所又は居所】	東京都千代田区霞が関 1 丁目 3 番 1 号
【氏名又は名称】	工業技術院長

【特許出願人】

【識別番号】	592163860
【住所又は居所】	東京都大田区南千束 1 丁目 9 番 13 号
【氏名又は名称】	後藤 真孝

【特許出願人】

【識別番号】	500331172
【住所又は居所】	茨城県つくば市梅園 1 丁目 1 番 4 工業技術院電 子技術総合研究所内

【氏名又は名称】	伊藤 克亘
----------	-------

【指定代理人】

申請人	
【識別番号】	220000356
【住所又は居所】	茨城県つくば市梅園 1 丁目 1 番 4
【氏名又は名称】	工業技術院電子技術総合研究所長

出 願 人 履 歴 情 報

識別番号 [000001144]

1. 変更年月日	1990年 9月20日
[変更理由]	新規登録
住 所	東京都千代田区霞が関1丁目3番1号
氏 名	工業技術院長

出 願 人 履 歴 情 報

識別番号 [592163860]

1. 変更年月日 1992年 7月 6日

[変更理由] 新規登録

住 所 東京都大田区南千束1丁目9番13号

氏 名 後藤 真孝

出 願 人 履 歴 情 報

識別番号 [500331172]

1. 変更年月日 2000年 7月13日

[変更理由] 新規登録

住 所 茨城県つくば市梅園1丁目1番4 工業技術院電子技術総合研  
究所内

氏 名 伊藤 克亘